(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 035 538 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
13.09.2000 Bulletin 2000/37

(51) Int. Cl.^{7}: **G10L 19/14**, G10L 19/12

(21) Application number: 00200874.6

(22) Date of filing: 13.03.2000

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 12.03.1999 US 124112 P
12.03.1999 US 124089 P

(71) Applicant:
Texas Instruments Incorporated
Dallas, TX 75251 (US)

(72) Inventors:
- **Stachurski, Jacek**
  Dallas, Texas 75243 (US)
- **McCree, Alan V.**
  Dallas Texas 75248 (US)

(74) Representative: **Holt, Michael**
Texas Instruments Limited,
European Patents Department (MS 13),
PO Box 5069
Northampton NN4 7ZE (GB)

(54) Multimode quantizing of the prediction residual in a speech coder

(57) Linear predictive system with classification of LP residual Fourier coefficients into two or more overlapping classes, and each class has its own vector quantization codebook(s). And modified use of strong and weak predictors to replace a strong predictor following a weak predictor with a weak predictor to insure attenuation of error propagation as arise from frame erasures.

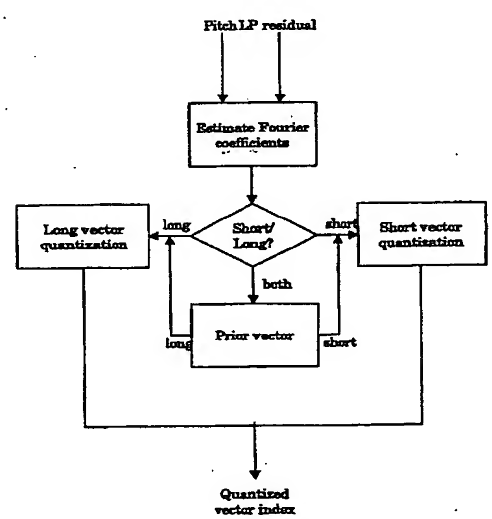

Figure 1a

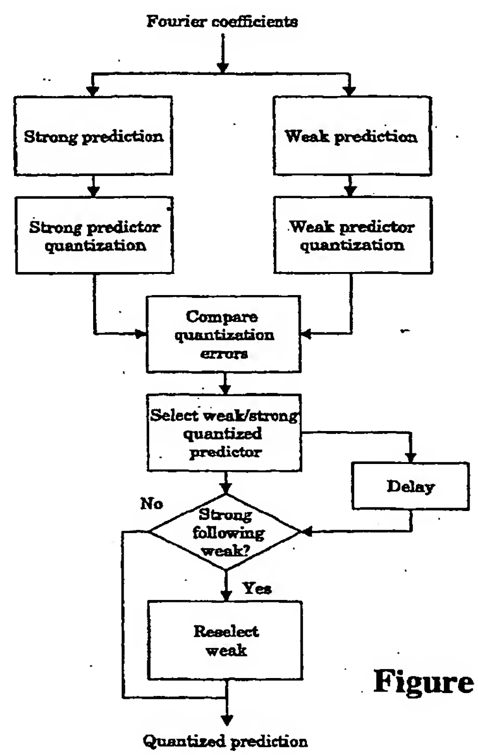

Figure 1b

EP 1 035 538 A2

Description

FIELD OF THE INVENTION

[0001] The present invention relates generally to the field of electronic devices, and, more particularly, to speech coding, technical transmission, storage, and synthesis circuitry and methods.

[0002] The performance of digital speech systems using low bits rates has become increasingly important with current and foreseeable digital communications. One digital speech method, linear predictive coding (LPC), uses a parametric model to mimic human speech. In this approach only the parameters of the speech model are transmitted across the communication channel (or stored), and a synthesizer regenerates the speech with the same perceptual characteristics as the input speech waveform. Periodic updating of the model parameters requires fewer bits than direct representation of the speech signal, so a reasonable LPC vocoder can operate at bits rates as low as 2-3 Kbps (kilobits per second) whereas the public telephone system uses 64 Kbps (8 bit PCM codewords at 8,000 samples per second). See for example, McCree et al, A 2.4 Kbit/s MELP Coder Candidate for the New U.S. Federal Standard, Proc. IEEE Int.Conf.ASSP 200 (1996) and US Patent No.5,699,477.

[0003] However, the speech output from such LPC vocoders is not acceptable in many applications because it does not always sound like natural human speech, especially in the presence of background noise. And there is a demand for a speech vocoder with at least telephone quality speech at a bit rate of about 4 Kbps. Various approaches to improve quality include enhancing the estimation of the parameters of a mixed excitation linear prediction (MELP) system and more efficient quantization of them. See Yeldener et al, A Mixed Sinusoidally Excited Linear Prediction coder at 4 kb/s and Below, Proc. IEEE Int. Conf. Acoust.,Speech,Signal Processing (1998) and Shlomot et al, Combined Harmonic and Waveform Coding of Speech at Low Bit Rates, IEEE ... 585 (1998).

SUMMARY OF THE INVENTION

[0004] The present application discloses a linear predictive coding method with the residual's Fourier coefficients classified into overlapping classes with each class having its own vector quantization codebook(s).

[0005] Additionally, both strongly predictive and weakly predictive codebooks may be used but with a weak predictor replacing a strong predictor which otherwise would have followed a weak predictor.

[0006] This has the advantages including maintenance of low bit rates but with increased performance and avoidance of error propagation by a series of strong predictors.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Specific embodiments of the present invention will now be described in further detail, by way of example, with reference to the accompanying drawings in which:

Figures 1a-1b are flow diagrams of a preferred embodiments.

Figures 2a-2b illustrate preferred embodiment coder and decoder in block format; and

Figures 3a-3d show an LP residual and its Fourier transforms.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0008] First preferred embodiments classify the spectra of the linear prediction (LP) residual (in a MELP coder) into classes of spectra (vectors) and vector quantize each class separately. For example, one first preferred embodiment classifies the spectra into long vectors (many harmonics which correspond roughly to low pitch frequency as typical of male speech) and short vectors (few harmonics which correspond roughly to high pitch frequency as typical of female speech). These spectra are then vector quantized with separate codebooks to facilitate encoding of vectors with different numbers of components (harmonics). Figure 1a shows the classification flow and includes an overlap of the classes.

[0009] Second preferred embodiments allow for predictive coding of the spectra (or alternatively, other parameters such as line spectral frequencies or LSFs) and a selection of either the strong or weak predictor based on best approximation but with the proviso that a first strong predictor which otherwise follows a weak predictor is replaced with a weak predictor. This deters error propagation by a sequence of strong predictors of an error in a weak predictor preceding the series of strong predictors. Figure 1b illustrates a predictive coding control flow.

[0010] Figures 2a-2b illustrate preferred embodiment MELP coding (analysis) and decoding (synthesis) in block format. In particular, the Linear Prediction Analysis determines the LPC coefficients a(j), j = 1, 2, ..., M, for an input frame of digital speech samples {y(n)} by setting:

$$e(n) = y(n) - \Sigma_{M \geq j \geq 1} a(j)y(n-j) \qquad (1)$$

and minimizing $\Sigma e(n)^2$. Typically, M, the order of the linear prediction filter, is taken to be about 10-12; the sampling rate to form the samples y(n) is taken to be 8000 Hz (the same as the public telephone network sampling for digital transmission); and the number of samples {y(n)} in a frame is often 160 (a 20 msec frame) or 180 (a 22.5 msec frame). A frame of samples may be gener-

ated by various windowing operations applied to the input speech samples. The name "linear prediction" arises from the interpretation of $e(n) = y(n) - \Sigma_{M \geq j \geq 1} a(j)y(n-j)$ as the error in predicting y(n) by the linear sum of preceding samples $\Sigma_{M \geq j \geq 1} a(j)y(n-j)$. Thus minimizing $\Sigma e(n)^2$ yields the {a(j)} which furnish the best linear prediction. The coefficients {a(j)} may be converted to LSFs for quantization and transmission.

[0011] The {e(n)} form the LP residual for the frame and ideally would be the excitation for the synthesis filter 1/A(z) where A(z) is the transfer function of equation (1). Of course, the LP residual is not available at the decoder; so the task of the encoder is to represent the LP residual so that the decoder can generate the LP excitation from the encoded parameters.

[0012] The Band-Pass Voicing for a frequency band of samples (typically two to five bands, such as 0-500 Hz, 500-1000 Hz, 1000-2000 Hz, 2000-3000 Hz, and 3000-4000 Hz) determines whether the LP excitation derived from the LP residual {e(n)} should be periodic (voiced) or white noise (unvoiced) for a particular band.

[0013] The Pitch Analysis determines the pitch period (smallest period in voiced frames) by low pass filtering {y(n)} and then correlating {y(n)} with {y(n+m)} for various m; interpolations provide for fractional sample intervals. The resultant pitch period is denoted pT where p is a real number, typically constrained to be in the range 20 to 132 and T is the sampling interval of 1/8 millisecond. Thus p is the number of samples in a pitch period. The LP residual {e(n)} in voiced bands should be a combination of pitch-frequency harmonics.

[0014] Fourier Coeff. Estimation provides coding of the LP residual for voiced bands. The following sections describe this in detail.

[0015] Gain Analysis sets the overall energy level for a frame.

[0016] The encoding (and decoding) may be implemented with a digital signal processor (DSP) such as the TMS320C30 manufactured by Texas Instruments which can be programmed to perform the analysis or synthesis essentially in real time.

[0017] Figure 3a illustrates an LP residual {e(n)} for a voiced frame and includes about eight pitch periods with each pitch period about 26 samples. Figure 3b shows the magnitudes of the {E(j)} for one particular period of the LP residual, and Figure 3c shows the magnitudes of the {E(j)} for all eight pitch periods. For a voiced frame with pitch period equal to pT, the Fourier coefficients peak about 1/pT, 2/pT, 3/pT, ..., k/pT, ...; that is, at the fundamental frequency 1/pT and harmonics. Of course, p may not be an integer, and the magnitudes of the Fourier coefficients at the fundamental-frequency harmonics, denoted X[1], X[2], ..., X[k], ... must be estimated. These estimates will be quantized, transmitted, and used by the decoder to create the LP excitation.

[0018] The {X[k]} may be estimated by various methods: for example, apply a discrete Fourier transform to the samples of a single period (or small number of periods) of e(n) as in Figures 3b-3c; alternatively, the {E(j)} can be interpolated. Indeed, one interpolation approach applies a 512-point discrete Fourier transform to an extended version of the LP residual, which allows use of a fast Fourier transform. In particular, extend the LP residual {e(n)} of 160 samples to 512 samples by setting $e_{512}(n) = e(n)$ for n = 0, 1, ... ,159, and $e_{512}(n) = 0$ for n = 160, 161, ..., 511. Then the discrete Fourier transform magnitudes appear as in Figure 3d with coefficients $E_{512}(j)$ which essentially interpolate the coefficients E(j) of Figures 3b-3c. Estimate the peaks X[k] at frequencies k/pT. The preferred embodiment only uses the magnitudes of the Fourier coefficients, although the phases could also be used. Because the LP residual components {e(n)} are real, the discrete Fourier transform coefficients {E(j)} are conjugate symmetric: $E(k) = E^*(N-k)$ for an N-point discrete Fourier transform. Thus only half of the {E(j)} need be used for magnitude considerations.

[0019] Once the estimated magnitudes of the Fourier coefficients X[k] for the fundamental pitch frequency and harmonics k/pT have been found, they must be transmitted with a minimal number of bits. The preferred embodiments use vector quantization of the spectra. That is, treat the set of Fourier coefficients X[1], X[2], ... X[k], ... as a vector in a multi-dimensional quantization, and transmit only the index of the output quantized vector. Note that there are [p] or [p]+1 coefficients, but only half of the components are significant due to their conjugate symmetry. Thus for a short pitch period such as pT = 4 milliseconds (p = 32), the fundamental frequency 1/pT (= 250 Hz) is high and there are 32 harmonics, but only 16 would be significant (not counting the DC component). Similarly, for a long pitch period such as pT = 12 milliseconds (p = 96), the fundamental frequency (= 83 Hz) is low and there are 48 significant harmonics.

[0020] In general, the set of output quantized vectors may be created by adaptive selection with a clustering method from a set of input training vectors. For example, a large number of randomly selected vectors (spectra) from various speakers can be used to form a codebook (or codebooks with multistep vector quantization) . Thus a quantized and coded version of an input spectrum X[1], X[2], ... X[k], ... can be transmitted as the index in the codebook of the quantized vector and which may be 20 bits.

[0021] As illustrated in Figure 1a, the first preferred embodiments proceed with vector quantization of the Fourier coefficient spectra as follows. First, classify a Fourier coefficient spectrum (vector) according to the corresponding pitch period: if the pitch period is less than 55T, the vector is a "short" vector, and if the pitch period is more than 45T, the vector is a "long" vector. Some vectors will qualify as both short and long vectors. Vector quantize the short vectors with a codebook of 20-component vectors, and vector quantize the long vectors with a codebook of 45-component vectors. As

described previously, conjugate symmetry of the Fourier coefficients implies only the first half of the vector components are significant and used. And for short vectors with less than 20 significant components, expand to 20 components by appending components equal to 1. Analogously for long vectors with fewer than 45 significant components, expand to 45 components by appending components equal to 1. Each codebook has 2^{20} output quantized vectors, so 20 bits will index the output quantized vectors in each codebook. One bit could be used to select the codebook, but the pitch is transmitted and can be used to determine whether the 20 bits are long or short vector quantization.

[0022] For a vector classified as both short and long, use the same classification as the preceding frame's vector; this avoids discontinuities and provides a hysteresis by the classification overlap. Further, if the preceding frame was unvoiced, then take the vector as short if the pitch period is less than 50T and long otherwise.

[0023] Apply a weighting factor to the metric defining distance between vectors. The distance is used both for the clustering of training vectors (which creates the codebook) and for the quantization of Fourier component vectors by minimum distance. In general, define a distance between vectors X_1 and X_2 by $d(X_1, X_2) = (X_1 - X_2)^T W (X_1 - X_2)$ with W a matrix of weights. Thus define matrices W_{short} for short vectors and matrices W_{long} for long vectors; further, the weights may depend upon the length of the vector to be quantized. Then for short vectors take $W_{short}[j,k]$ very small for either j or k larger than 20; this will render the components $X_1[k]$ and $X_2[k]$ irrelevant for k larger than 20. Further, take $W_{short}[j,k]$ decreasing as j and k increase from 1 to 20 to emphasize the lower vector components. That is, the quantization will depend primarily upon the Fourier coefficients for the fundamental and low harmonics of the pitch frequency. Analogously, take $W_{long}[j,k]$ very small for j or k larger than 45.

[0024] Further, the use of predictive coding could be included to reduce the magnitudes and decrease the quantization noise as described in the following.

Predictive coding

[0025] A differential (predictive) approach will decrease the quantization noise. That is, rather than vector quantize a spectrum X[1], X[2], ... X[k], ..., first generate a prediction of the spectrum from the preceding one or more frames' quantized spectra (vectors) and just quantize the difference. If the current frame's vector can be well approximated from the prior frames' vectors, then a "strong" prediction can be used in which the difference between the current frame's vector and a strong predictor may be small. Contrarily, if the current frame's vector cannot be well approximated from the prior frames' vectors, then a "weak" prediction (including no prediction) can be used in which the difference between the current frame's vector and a predictor may be large. For example, a simple prediction of the current frame's vector X could be the preceding frame's quantized vector Y, or more generally a multiple αY with α a weight factor (between 0 and 1). Indeed, α could be a diagonal matrix with different factors for different vector components. For α values in the range 0.7-1.0, the predictor αY is close to Y and if also close to X, the difference vector $X-\alpha Y$ to be quantized is small compared to X. This would be a strong predictor, and the decoder recovers an estimate for X by $Q(X-\alpha Y) + \alpha Y$ with the first term the quantized difference vector $X-\alpha Y$ and the second term from the previous frame and likely the dominant term. Conversely, for α values in the range 0.0-0.3, the predictor is weak in that the difference vector $X-\alpha Y$ to be quantized is likely comparable to X. In fact, $\alpha = 0$ is no prediction at all and the vector to be quantized is X itself.

[0026] The advantage of strong predictors follows from the fact that with the same size codebooks, quantizing something likely to be small (strong-predictor difference) will give better average results than quantizing something likely to be large (weak-predictor difference).

[0027] Thus train four codebooks: (1) short vectors and strong prediction, (2) short vectors and weak prediction, (3) long vectors and strong prediction, and (4) long vectors and weak prediction. Then process a vector as illustrated in the top portion of Figure 1b: first the vector X is classified as short or long; next, the strong and weak predictor vectors, X_{strong} and X_{weak}, are generated from previous frames' quantized vectors and the strong predictor and weak predictor codebooks are used for vector quantization of $X-X_{strong}$ and $X-X_{weak}$, respectively. Then the two results ($Q(X-X_{strong}) + X_{strong}$ and $Q(X-X_{weak}) + X_{weak}$) are compared to the input vector and the better approximation (strong or weak predictor) is selected. A bit is transmitted (to indicate whether a strong or weak predictor was used) along with the 20-bit codebook index for the quantization vector. The pitch determines whether the vector was long or short.

[0028] In a frame erasure the parameters (i.e., LSFs, Fourier coefficients, pitch, ...) corresponding to the current frame are considered lost or unreliable and the frame is reconstructed based on the parameters from the previous frames. In the presence of frame erasures the error resulting from missing a set of parameters will propagate throughout the series of frames for which a strong prediction is used. If the error occurs in the middle of the series, the exact evolution of the predicted parameters is compromised and some perceptual distortion is usually introduced. When a frame erasure happens within a region where a weak predictor is consistently selected, the effect of the error will be localized (it will be quickly reduced by the weak prediction). The largest degradation in the reconstructed frame is observed whenever a frame erasure occurs for a frame with a weak predictor followed by a

series of frames for which a strong predictor is chosen. In this case the evolution of the parameters is builtup on a parameter very different from that which is supposed to start the evolution.

[0029] Thus a second preferred embodiment analyzes the predictors used in a series of frames and controls their sequencing. In particular, for a current frame which otherwise would use a strong predictor immediately following a frame which used a weak predictor, one preferred embodiment modifies the current frame to use the weak predictor but does not affect the next frame's predictor. Figure 1b illustrates the decisions.

[0030] A simple example will illustrate the effect of this preferred embodiment. Presume a sequence of frames with Fourier coefficient vectors X_1, X_2, X_3, ... and presume the first frame uses a weak predictor and the second, third, fourth, ... frames use strong predictors, but the preferred embodiment replaces the second frame's strong predictor with a weak predictor. Thus the transmitted quantized difference vector for the first frame is $Q(X_1\text{-}X_{1weak})$ and without erasure the decoder recovers X_1 as $Q(X_1\text{-}X_{1weak}) + X_{1weak}$ with the first term likely the dominant term due to weak prediction. Similarly, the usual decoder recovers X_2 as $Q(X_2\text{-}X_{2strong}) + X_{2strong}$ with the second term dominant, and analogously for X_3, X_4, ... In contrast, the preferred embodiment decoder recovers X_2 as $Q(X_2\text{-}X_{2weak}) + X_{2weak}$ but with the first term likely dominant.

[0031] Note that the decoder recreates X_{1weak} from the preceding reconstructed frames' vectors X_0, X_1, ... , and similarly for $X_{2strong}$ and X_{2weak} recreated from reconstructed X_1, X_0, ..., and likewise for the other predictors.

[0032] Now with an erasure of the first frame parameters the vector $Q(X_1\text{-}X_{1weak})$ is lost and the decoder reconstructs the X_1 by something such as just repeating reconstructed X_0 from the prior frame. However, this may not be a very good approximation because originally a weak predictor was used. Then for the second frame, the usual decoder reconstructs X_2 by $Q(X_2\text{-}X_{2strong}) + Y_{2strong}$ with $Y_{2strong}$ the strong predictor recreated from X_0, X_0, ... rather than from X_1, X_0, ... because X_1 was lost and replaced by possibly poor approximation X_0. Thus the error would roughly be $X_{2strong} - Y_{2strong}$ which likely is large due to the strong predictor being the dominant term compared to the difference term $Q(X_2\text{-}X_{2strong})$. And this also applies to the reconstruction of X_3, X_4,...

[0033] Contrarily, the preferred embodiment reconstructs X_2 by $Q(X_2\text{-}X_{2weak}) + Y_{2weak}$ with $Y_{2strong}$ the weak predictor recreated from X_0, X_0, ... rather than from X_1, X_0, ... again because X_1 was lost and replaced by possibly poor approximation X_0. Thus the error would roughly be $X_{2weak} - Y_{2weak}$ which likely is small due to the weak predictor being the smaller term compared to the difference term $Q(X_2\text{-}X_{2weak})$. And this smaller error also applies to the reconstruction of X_3, X_4,

[0034] Indeed for the case of the predictors $X_{2strong} = \alpha X_1$ with $\alpha = 0.8$ and $X_{2weak} = \alpha X_1$ with $\alpha = 0.2$, the usual decoder error would be $0.8(X_1 - X_0)$ for reconstruction of X_2 and the preferred embodiment decoder error would be $0.2(X_1 - X_0)$.

[0035] Alternative second preferred embodiments modify two (or more) successive frame's strong predictors after a weak predictor frame to be weak predictors. That is, a sequence of weak, strong, strong, strong, ... would be changed to weak, weak, weak, strong, ...

[0036] The foregoing replacement of strong predictors by weak predictors provides a tradeoff of increased error robustness for slightly decreased quality (the weak predictors being used in place of better strong predictors).

Claims

1. A method of Linear predictive system coding, comprising the steps of:

   classifying LP residual Fourier coefficients into two or more classes of vectors;
   for each class of vectors providing at least one vector quantization codebook; and
   encoding said vectors with said codebooks.

2. The coding of claim 1, wherein said classes of vectors overlap and a vector in two or more classes is encoded using the class of a vector in a preceding frame.

3. A method of Linear predictive system decoding, comprising the steps of:

   interpreting LP residual Fourier coefficients as members of two or more overlapping classes of vectors with each class having at least one vector quantization codebook; and
   decoding an encoded vector using said codebooks.

4. An encoding method using strong and weak predictors, comprising the step of:

   replacing a strong predictor following a weak predictor with a weak predictor.

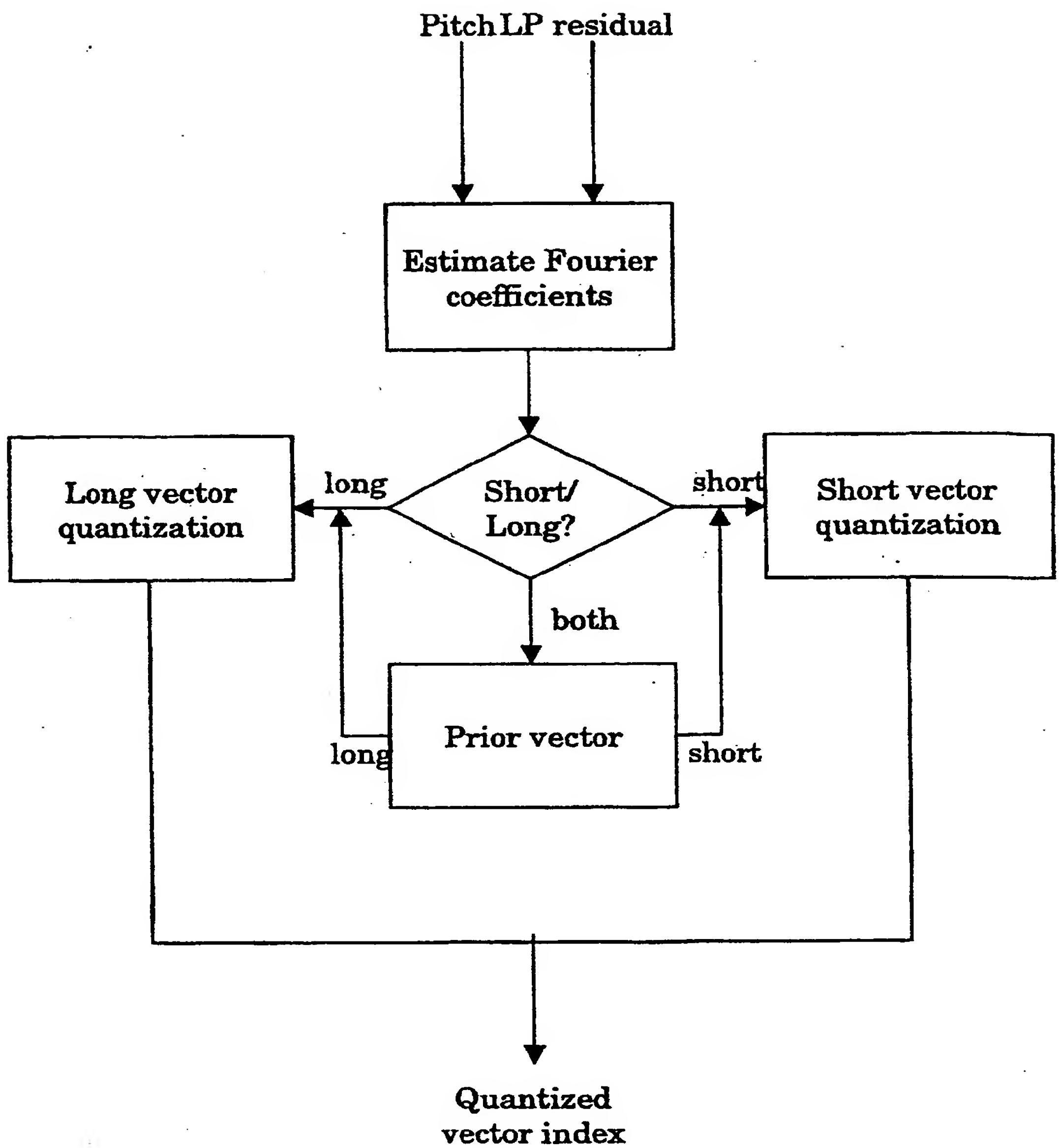
Pitch LP residual
Estimate Fourier coefficients
Short/ Long?
long
short
both
Long vector quantization
Short vector quantization
Prior vector
long
short
Quantized vector index

Figure 1a

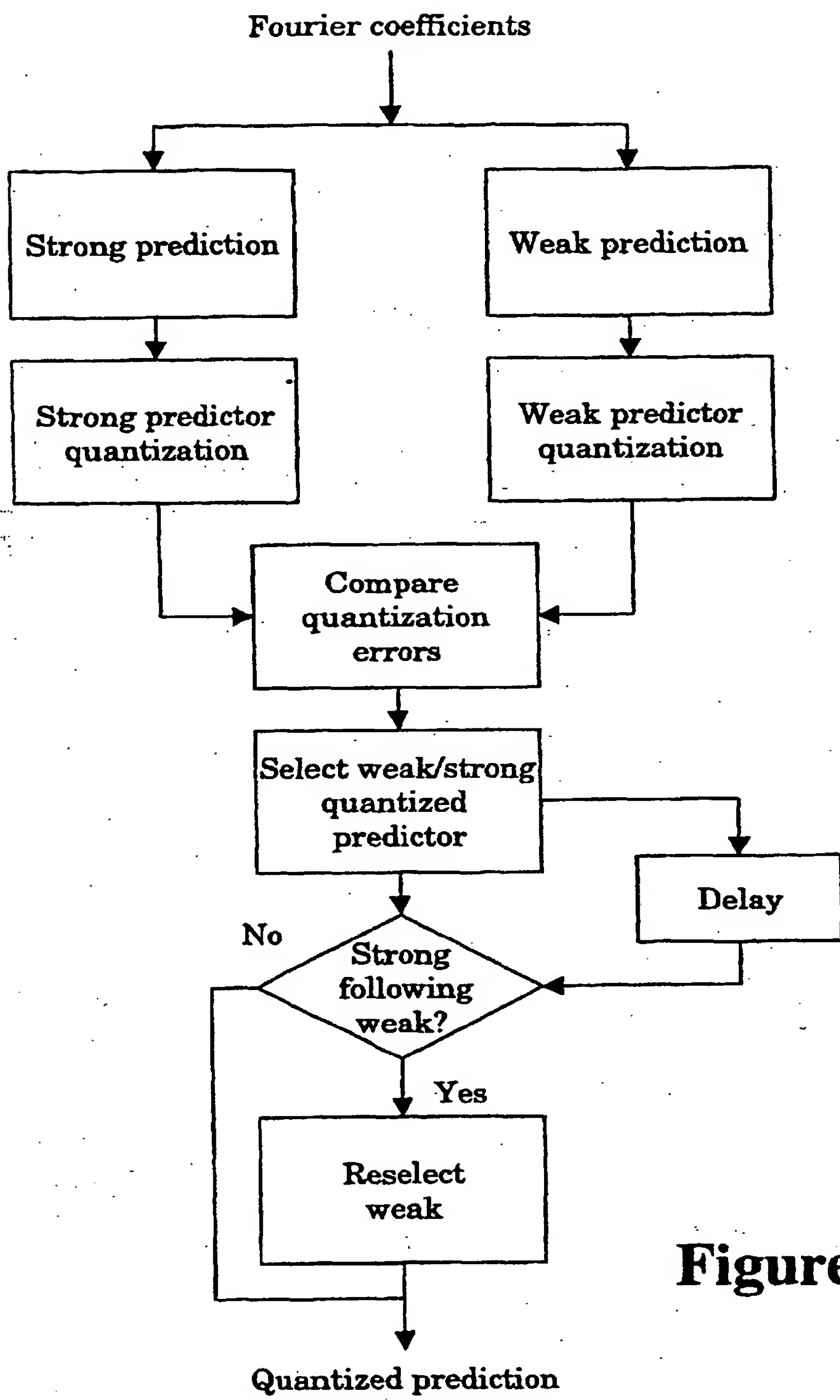
Fourier coefficients
Strong prediction
Weak prediction
Strong predictor quantization
Weak predictor quantization
Compare quantization errors
Select weak/strong quantized predictor
Delay
No
Strong following weak?
Yes
Reselect weak
Quantized prediction

Figure 1b

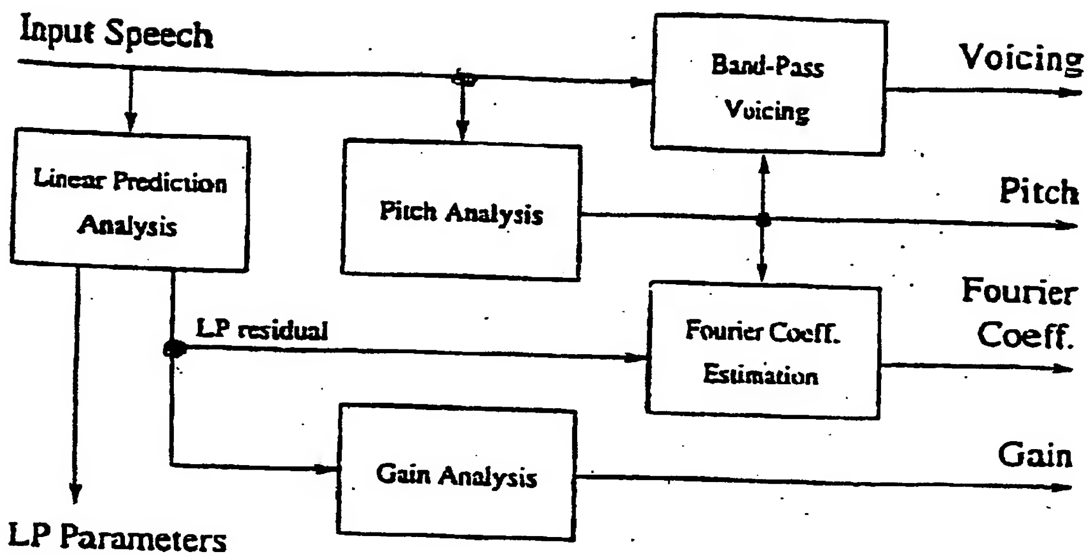

Figure 2a MELP analysis

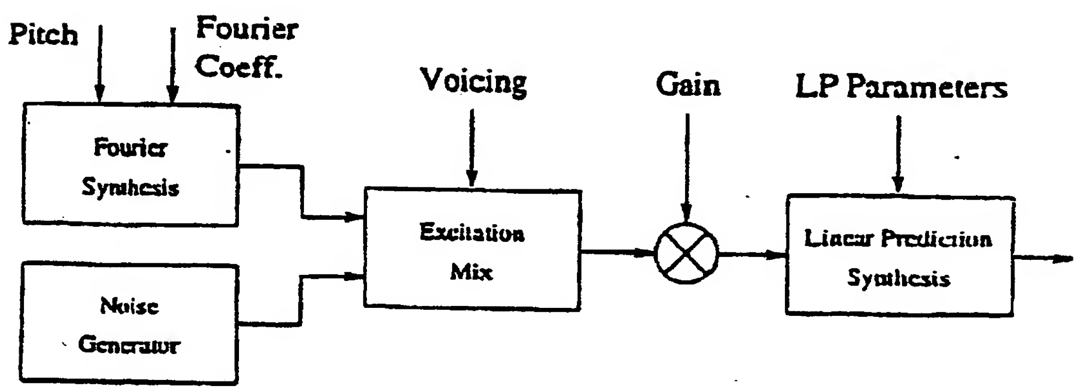

Figure 2b MELP synthesis

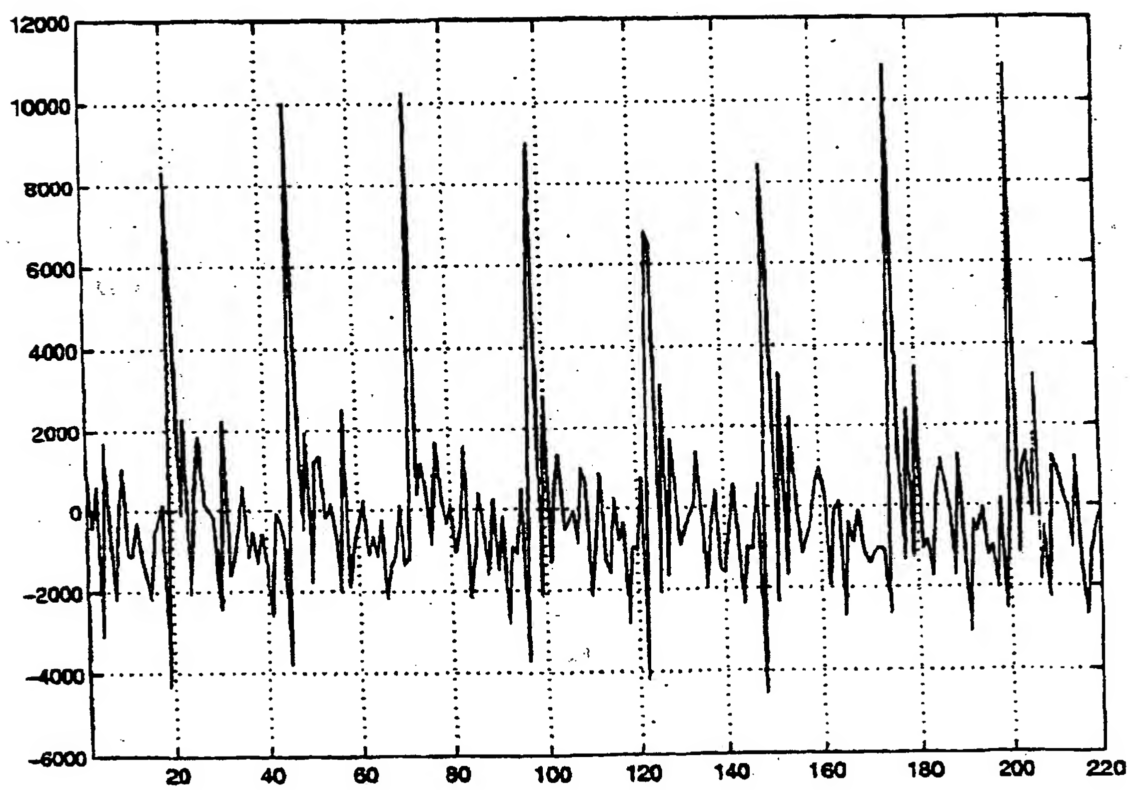
12000
10000
8000
6000
4000
2000
0
-2000
-4000
-6000
20
40
60
80
100
120
140
160
180
200
220

Fig 3a

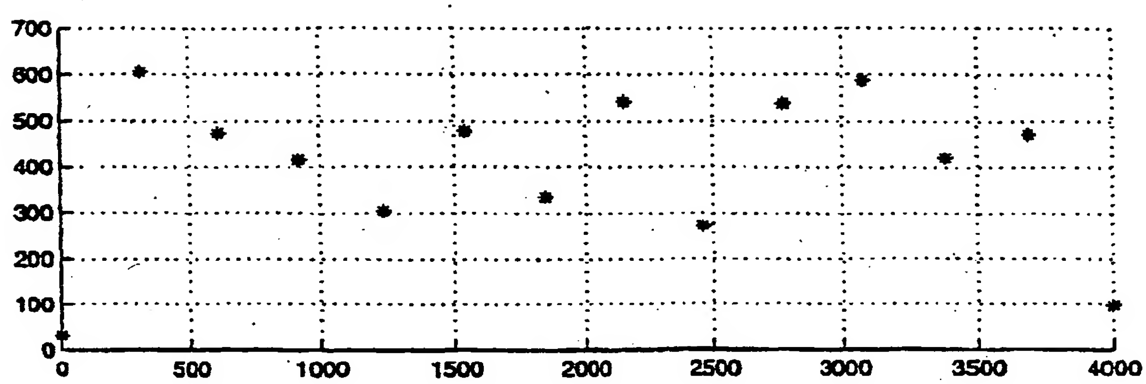
700
600
500
400
300
200
100
0
0
500
1000
1500
2000
2500
3000
3500
4000

Fig 3b

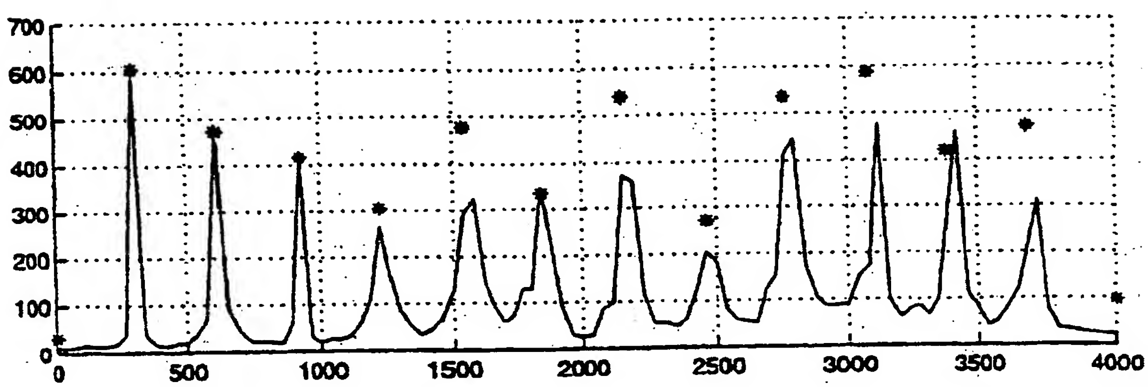
700
600
500
400
300
200
100
0
0
500
1000
1500
2000
2500
3000
3500
4000

Fig 3c

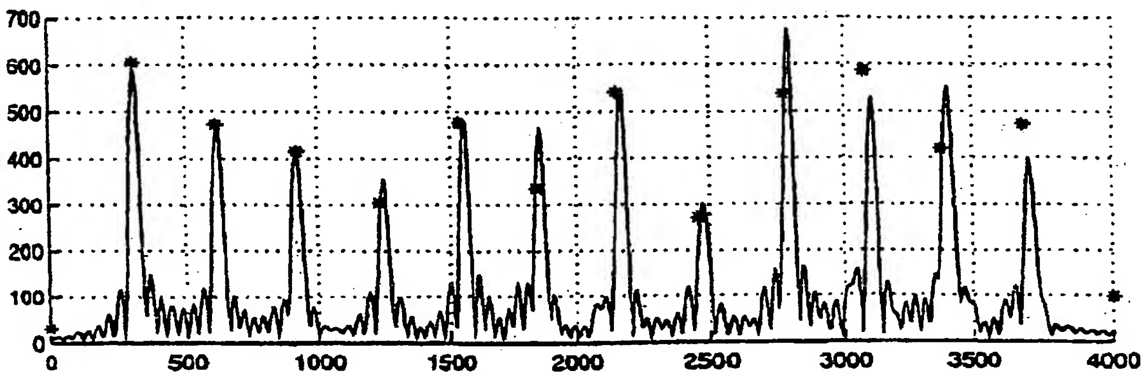
700
600
500
400
300
200
100
0
0
500
1000
1500
2000
2500
3000
3500
4000

Fig 3d

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 035 538 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
23.04.2003 Bulletin 2003/17

(43) Date of publication A2:
13.09.2000 Bulletin 2000/37

(21) Application number: 00200874.6

(22) Date of filing: 13.03.2000

(51) Int Cl.7: G10L 19/14, G10L 19/12

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 12.03.1999 US 124112 P
12.03.1999 US 124089 P

(71) Applicant: Texas Instruments Incorporated
Dallas, TX 75251 (US)

(72) Inventors:
- Stachurski, Jacek
  Dallas, Texas 75243 (US)
- McCree, Alan V.
  Dallas Texas 75248 (US)

(74) Representative: Holt, Michael
Texas Instruments Limited,
European Patents Department (MS 13),
PO Box 5069
Northampton NN4 7ZE (GB)

(54) Multimode quantizing of the prediction residual in a speech coder

(57) Linear predictive system with classification of LP residual Fourier coefficients into two or more overlapping classes, and each class has its own vector quantization codebook(s). And modified use of strong and weak predictors to replace a strong predictor following a weak predictor with a weak predictor to insure attenuation of error propagation as arise from frame erasures.

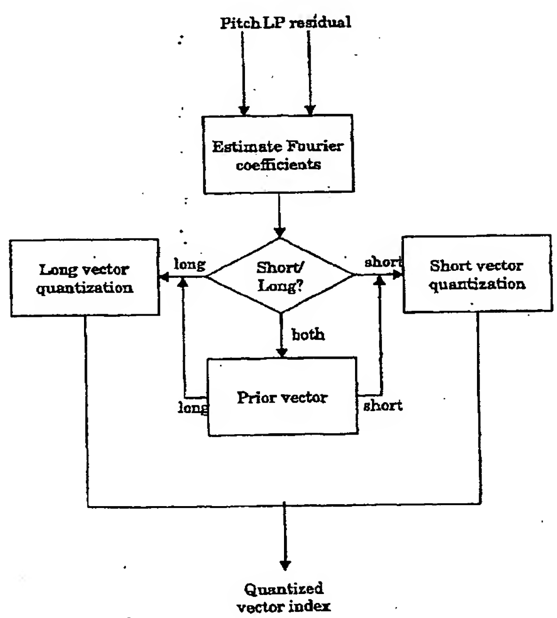

Figure 1a

Printed by Jouve, 75001 PARIS (FR)

(Cont. next page)

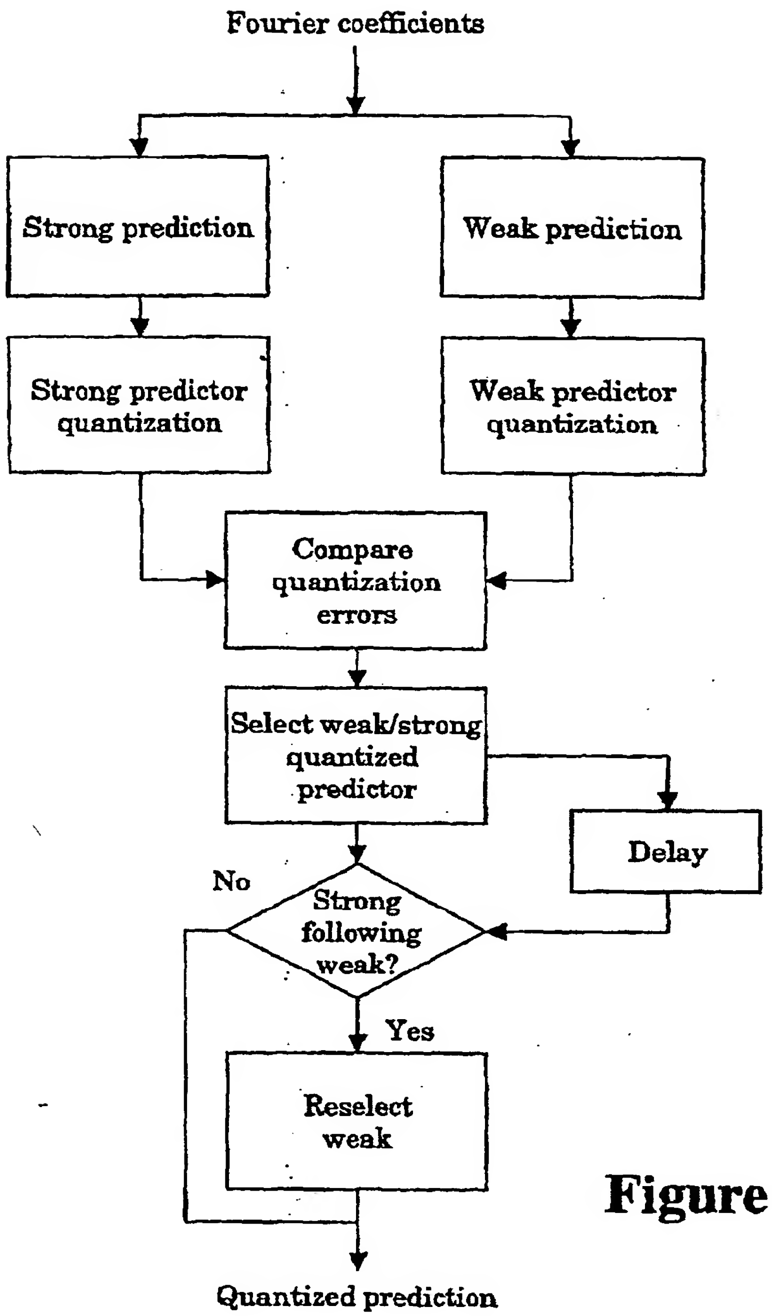
Fourier coefficients
Strong prediction
Weak prediction
Strong predictor quantization
Weak predictor quantization
Compare quantization errors
Select weak/strong quantized predictor
Delay
No
Strong following weak?
Yes
Reselect weak
Quantized prediction

Figure 1b

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 20 0874

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X A	US 5 749 065 A (NISHIGUCHI MASAYUKI ET AL) 5 May 1998 (1998-05-05) * abstract; figure 9 * * column 3, line 52 - line 62 * * column 16, line 37 - column 17, line 28 *	1,3 2	G10L19/14 G10L19/12
A	EP 0 751 494 A (SONY CORP) 2 January 1997 (1997-01-02) * abstract; figure 1 * * column 5, line 32 - column 6, line 52 *	1-3	
A	MARSTON D F: "Gender adapted speech coding" ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 1998. PROCEEDINGS OF THE 1998 IEEE INTERNATIONAL CONFERENCE ON SEATTLE, WA, USA 12-15 MAY 1998, NEW YORK, NY, USA,IEEE, US, 12 May 1998 (1998-05-12), pages 357-360, XP010279165 ISBN: 0-7803-4428-6 * abstract *	1-3	TECHNICAL FIELDS SEARCHED (Int.Cl.7) G10L

-/--

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	11 February 2003	Zimmermann, E

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent
Office

Application Number
EP 00 20 0874

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

☐ Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid, namely claim(s):

☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

☒ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.

☐ As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.

☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:

☐ None of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 20 0874

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
P,X	STACHURSKI JACEK ET AL: "High quality MELP coding at bit-rates around 4 kb/s" PROCEEDINGS OF THE 1999 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP-99);PHOENIX, AZ, USA MAR 15-MAR 19 1999, vol. 1, 1999, pages 485-488, XP002230714 ICASSP IEEE Int Conf Acoust Speech Signal Process Proc;ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings 1999 IEEE, Piscataway, NJ, USA * page 487, left-hand column, line 48 - right-hand column, line 19 * ---	4	
A	ERIKSSON T ET AL: "Exploiting interframe correlation in spectral quantization: a study of different memory VQ schemes" 1996 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING CONFERENCE PROCEEDINGS (CAT. NO.96CH35903), 1996 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING CONFERENCE PROCEEDINGS, ATLANTA, GA, USA, 7-10 M, pages 765-768 vol. 2, XP002230715 1996, New York, NY, USA, IEEE, USA ISBN: 0-7803-3192-3 * page 765, right-hand column, line 26 - page 766, right-hand column, line 31; figure 2 * ---	4	TECHNICAL FIELDS SEARCHED (Int.Cl.7)

-/--

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	11 February 2003	Zimmermann, E

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

LACK OF UNITY OF INVENTION SHEET B

Application Number
EP 00 20 0874

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. Claims: 1-3

   Classifying residual coefficients into two or more classes of vectors and encoding the vectors using codebooks that pertain to the selected class.

2. Claim : 4

   Replacing a strong predictor following a weak predictor with a weak predictor.

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 20 0874

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	MCCREE A ET AL: "A 1.7 kb/s MELP coder with improved analysis and quantization" ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 1998. PROCEEDINGS OF THE 1998 IEEE INTERNATIONAL CONFERENCE ON SEATTLE, WA, USA 12-15 MAY 1998, NEW YORK, NY, USA,IEEE, US, 12 May 1998 (1998-05-12), pages 593-596, XP010279196 ISBN: 0-7803-4428-6 * page 594, right-hand column, line 33 - page 595, left-hand column, line 4; figure 2 * -----	4	
			TECHNICAL FIELDS SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	11 February 2003	Zimmermann, E

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 00 20 0874

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

11-02-2003

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5749065 A	05-05-1998	JP 8069299 A	12-03-1996
EP 0751494 A	02-01-1997	JP 8179796 A	12-07-1996
		AU 703046 B2	11-03-1999
		AU 4190196 A	10-07-1996
		BR 9506841 A	14-10-1997
		EP 0751494 A1	02-01-1997
		PL 316008 A1	23-12-1996
		US 5950155 A	07-09-1999
		CA 2182790 A1	27-06-1996
		CN 1141684 A	29-01-1997
		WO 9619798 A1	27-06-1996
		TR 960601 A2	21-07-1996

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82